J. Mol. Biol. (1974) **89**, 179–193

Deletion Mutants of Simian Virus 40 Generated by Enzymatic Excision of DNA Segments from the Viral Genome

CHING-JUH LAI AND DANIEL NATHANS

Department of Microbiology
The Johns Hopkins University School of Medicine
Baltimore, Md 21205, U.S.A.

(Received 9 April 1974)

Deleted genomes of simian virus 40 have been constructed by enzymatic excision of specific segments of DNA from the genome of wild-type SV40†. For this purpose, a restriction endonuclease from *Hemophilus influenzae* (endo R·*Hin*dIII) was used. This enzyme cleaves SV40 DNA into six fragments, which have cohesive termini. Partial digest products were separated by electrophoresis in agarose gel and subsequently cloned by plaque formation in the presence of complementing temperature-sensitive mutants of SV40. Individual deletion mutants generated in this way were mapped by analysis of DNA fragments produced by endo R·*Hin*d digestion of their deleted genomes, and by heteroduplex mapping. Two types of deletions were found: (1) "excisional" deletions, in which the limits of the deleted segment corresponded to *Hin*dIII cleavage sites, and (2) "extended" deletions, in which the deleted segment extended beyond *Hin*dIII cleavage sites. Excisionally deleted genomes presumably arose by cyclization of a linear fragment *via* cohesive termini generated by endo R·*Hin*dIII whereas genomes with extended deletions probably were generated by intramolecular recombination near the ends of linear fragments. Of the nine mutants analyzed, two had deletions in the "early" region of the SV40 genome, six had deletions in the "late" region, and one had a deletion that spanned both regions.

1. Introduction

As one of the simplest oncogenic viruses, simian virus 40 is under intensive investigation as a model for viral tumorigenesis and regulation of gene expression in animal cells. The DNA genome of SV40† is a circular duplex of about 5500 nucleotide pairs. Four complementation groups have been distinguished using temperature-sensitive mutants of the virus (Tegtmeyer & Ozer, 1971; Kimura & Dulbecco, 1972,1973; Robb *et al.*, 1974; Dubbs *et al.*, 1974; Chou & Martin, 1974); and the overall pattern of organization and expression of the SV40 genome is becoming clearer as a result of studies on SV40 transcription (Khoury *et al.*, 1972,1973; Sambrook *et al.*, 1972, 1973), DNA replication (Danna & Nathans, 1972; Fareed *et al.*, 1972), mapping of adeno–SV40 hybrid viruses (Kelly & Lewis, 1973; Morrow *et al.*, 1973; Lebowitz *et al.*, 1974), and localization of ts mutations (Lai & Nathans, 1974).

† Abbreviations used: SV40, simian virus 40; DNA form I, closed circular SV40 DNA; DNA form II, open circular SV40 DNA; endo R·*Hin*d, *Hpa*, *Eco*, restriction endonucleases from *H. influenzae* strain d, *H. parainfluenzae* and *Escherichia coli*, respectively; *dl*, deletion mutant; ts, temperature-sensitive mutants.

To extend the genetic and physiological analysis of SV40 it would be of value to have deletion mutants of the virus, covering all regions of the genome. One source of such mutants are the light virions present in high multiplicity passage stocks of SV40 (Yoshiike, 1968; Tai *et al.*, 1972; Brockman *et al.*, 1973; Rozenblatt *et al.*, 1973). Brockman & Nathans (1974) have recently described a general method for cloning these defective variants. This method depends on the helper function of a ts mutant or wild-type virus to plaque-purify specific variants, and subsequent physical separation of mutant virus or viral DNA from the helper. Using this procedure, they have cloned complementing variants, as well as substituted variants, containing host cell DNA (Lavi & Winocour, 1972). Many of these variants contain duplications or rearrangements of genetic segments in addition to deletions (Tai *et al.*, 1972; Brockman *et al.*, 1973; Rozenblatt *et al.*, 1973; Martin *et al.*, 1974; Khoury *et al.*, 1974; Brockman & Nathans, 1974; Mertz & Berg, personal communication), and therefore are not ideal for certain physiological studies. Moreover, in selecting such "evolutionary" variants, there is no way to control the locus of alteration in the genome.

In order to obtain simple deletion mutants of SV40 involving different regions of the viral DNA, we have constructed deleted genomes by excising specific segments from wild-type DNA using restriction endonucleases which make several scissions in the molecule (Danna & Nathans, 1971). Such excisionally deleted molecules have then been cloned by complementation plaque formation in the presence of ts mutants (Brockman & Nathans, 1974). In this communication we describe the generation and analysis of several deletion mutants of this type, documenting that they contain simple deletions, and mapping the deletions with respect to restriction endonuclease cleavage sites. Although some of the deletions terminate at enzyme cleavage sites (excisional deletions), others extend beyond these sites (extended deletions), suggesting that they are formed by intramolecular recombination. In a subsequent communication we shall present the results of complementation tests and physiological studies with these mutants.

2. Materials and Methods

(a) *Cells and viruses*

The source of SV40 strains and African green monkey kidney (BSC) cells has been given in earlier publications (Danna & Nathans, 1971; Brockman & Nathans, 1974) as has the composition of media used. We are grateful to Dr Peter Tegtmeyer for supplying ts mutants of SV40 (Tegtmeyer & Ozer, 1971; Tegtmeyer, 1972).

(b) *Complementation plaqueing of SV40 DNA fragments with ts mutants*

This was carried out by first co-infecting BSC-40 cells in microwells and then plating infective centers, essentially as described earlier (Brockman & Nathans, 1974). Infection with DNA was done with DEAE-dextran as described by McCutchan & Pagano (1968). About 5×10^4 cells growing in microwells were exposed to 1 to 10 ng DNA in 25 μl of a solution consisting of minimal essential medium (Microbiological Associates), 0·05 M-Tris · HCl (pH 7·6), and 250 μg DEAE-dextran/ml. After incubation for 30 min at room temperature, the DNA solution was removed and the cells rinsed with minimal essential medium containing 2% fetal calf serum. The cells were then infected with an "early" ts mutant (tsA28) or with a "late" ts mutant (tsB4) at a multiplicity of 5 plaque-forming units cell. After incubation for 16 h at 37°C, the cells were transferred to 6-cm Petri dishes and allowed to attach at 37°C. Approximately 5×10^5 BSC-40 cells were then added and the dishes incubated for 16 to 20 h at 37°C. At this time the medium was replaced with 5 ml

of agar overlay and the dishes shifted to 40°C. After 8 days at 40°C agar containing 0·5% neutral red was added, and plaques were selected 2 days later.

(c) *Stocks*

Deletion mutants (together with helper ts mutant) were prepared from virus present in a single plaque as described earlier (Brockman & Nathans, 1974).

SV40 DNA was prepared by the procedure of Hirt (1967) as detailed previously (Danna *et al.*, 1973), using BSC-1 cells infected at 37°C with stocks of plaqued virus. Short viral genomes were separated by electrophoresis in 1·4% agarose (Brockman & Nathans, 1974).

(d) *Endonuclease R·*Hind

Endo R·*Hin*d was prepared by the procedure of Smith & Wilcox (1970) as modified by Smith (1974), except that the phosphocellulose column was developed by gradient elution with NaCl. To 50 ml of packed phosphocellulose, 200 ml of desalted ammonium sulfate fraction (from 100 g of cells originally) containing 12 mg protein/ml was applied overnight. The column was then washed sequentially with 100 ml of buffer (0·01 M-sodium phosphate (pH 7·4), 2 mM-mercaptoethanol) and 250 ml of buffer with 0·15 M-NaCl, followed by a 500-ml linear gradient of NaCl in buffer, 0·15 M to 0·35 M. Fractions of 10 ml were collected and tested for endonuclease activity by the DNA clot assay of Smith (1974). Active fractions were then tested by digestion of SV40 DNA and electrophoresis of products (Danna *et al.*, 1973). Endo R eluted at about 0·20 M to 0·24 M-NaCl. Pooled fractions were stored at −15°C in 15% glycerol.

(e) *Endonuclease R·*Hind*III*

One of the components of endo R·*Hin*d (Roy & Smith, 1973; Smith, 1974) was prepared from the above endo R·*Hin*d preparation. Assay by digestion of ^{32}P-labeled SV40 DNA and electrophoresis of products indicated that endo R·*Hin*d fractions eluting late from phosphocellulose were enriched for the dIII enzyme, which makes 6 scissions in SV40 DNA (Danna *et al.*, 1973; T. N. H. Lee, unpublished observations). These fractions were pooled, made 25% in glycerol, desalted on G25 Sephadex (equilibrated with 0·01 M-sodium phosphate (pH 7·4), 2 mM-mercaptoethanol, 25% glycerol), and applied to a 10-ml column of Whatman DE52 DEAE-cellulose equilibrated with the same buffer. The column was washed with 30 ml of starting buffer and then with 300 ml of a linear 0 M to 0·10 M-NaCl gradient in buffer. Endo R·*Hin*dIII was present only in the "run through" fractions and the initial wash. A second activity (endo R·*Hin*dII; Smith, 1974) was detected in fractions eluting at about 0·08 M-NaCl. Active fractions were stored at −15°C in 50% glycerol.

(f) *Digestion of SV40 DNA* with endo *R·*Hind*III*

Digestion was carried out under conditions previously described for endo R·*Hin*d (Danna *et al.*, 1973). When limit digests were to be prepared, preliminary experiments with varying amounts of enzyme were carried out to establish conditions for complete digestion as monitored by electrophoresis of fragments (see below). For partial enzymatic digestion, the amount of enzyme was reduced to give the desired products on electrophoresis. Cleavage of SV40 DNA with endo R·*Eco* RI (Hedgpeth *et al.*, 1972) was carried out at 37°C for 1 h in an incubation mixture containing 0·1 M-Tris·HCl (pH 7·6), 0·01 M-$MgCl_2$ and 4 μg DNA in a volume of 0·2 ml. (We are grateful to Dr Paul Geshelin for preparing this enzyme.) Digestion of SV40 DNA or of *Hin*dIII fragments with endo R·*Hin*d was carried out as previously described (Danna *et al.*, 1973), using varying amounts of enzyme to assure complete digestion.

(g) *Electrophoresis*

Electrophoresis of SV40 DNA or large *Hin*dIII fragments of SV40 DNA was performed in 1·4% agarose gel slabs as previously described (Brockman & Nathans, 1974). Fragments were eluted by electrophoresis and dialyzed against 15 mM-NaCl, 1·5 mM-sodium citrate (pH 7·4). *Hin* fragments of SV40 DNA were separated by electrophoresis in 4% polyacrylamide gel slabs (Danna *et al.*, 1973).

(h) *Electron microscopy of DNA*

This was done using the formamide method of Davis *et al.* (1971). For length measurements, a minimum of 40 DNA molecules were measured and compared to full-length SV40 DNA on the same grid. Heteroduplex DNA molecules composed of deleted and full-length strands were prepared by mixing in a total volume of 20 μl, 0·05 μg *Eco* RI SV40 linear DNA and 0·05 μg *Eco* RI deleted linear DNA in 0·1 M-NaOH. After 10 min at room temperature, 2 μl of 1 M-Tris · HCl (pH 7·1) were added. Renaturation was carried out by adding 22 μl of 98% formamide and incubating the mixture at room temperature for 1 h. The reannealed DNA solution containing 0·1 mg cytochrome *c*/ml, 40% formamide, and 0·01 M-Tris·HCl (pH 8·6) was spread on a 10% formamide hypophase, buffered with 0·01 M-Tris·HCl, pH 8·6. Specimens were stained with uranyl acetate, shadowed with platinum/palladium, and examined in an electron microscope.

3. Results

(a) *Experimental design*

The general approach used to construct and clone deleted genomes of SV40 was first to digest intact SV40 DNA form I with limiting concentrations of a restriction endonuclease which produces multiple scissions in the molecule. The partial digest products were then fractionated according to length by electrophoresis in agarose gel. Fragments ranging from about 65% to nearly 100% of the length of SV40 DNA were isolated and used to infect cells, which were then co-infected with an early or a late ts mutant. Infected cells were plated at non-permissive temperature and resulting plaques were tested for the presence of viruses with short genomes (Brockman & Nathans, 1974). The cloned, deleted genomes isolated by electrophoresis in agarose were then mapped by analyzing fragments produced by endo R·*Hin*d and by heteroduplex mapping.

In the studies to be reported we used restriction endonuclease *Hin*dIII (see below), which produces DNA fragments with cohesive ends (Old, Murray & Roizes, personal communication). Such fragments should form covalently closed circles in infected cells, as inferred for full-length linear SV40 DNA molecules with cohesive ends (Mertz & Davis, 1972). Subsequently, it was found that the presence of cohesive ends is not required to obtain deletion mutants by this procedure (see Discussion).

(b) *Cleavage of SV40 DNA by endo R·HindIII*

In a previous report, Danna *et al.* (1973) indicated that the restriction endonuclease from *Hemophilus influenzae* described by Smith & Wilcox (1970) contains two main enzymes which cleave SV40 DNA, designated endo R·*Hin*dII and endo R·*Hin*dIII. *Hin*dIII breaks SV40 DNA into six fragments. To determine the sites of cleavage, the dIII fragments were digested with the composite *Hin* enzyme (dII and dIII) and the products analyzed along the lines previously used to construct the *Hin* cleavage map (Danna *et al.*, 1973). Plate I shows the electrophoretic separation of the six *Hin*dIII fragments of SV40 DNA. Based on electrophoretic mobility relative to *Hin* fragments (Danna *et al.*, 1973) we estimate their lengths as 33% (dIII-A), 22·5% (dIII-B), 21% (dIII-C), 11% (dIII-D), 9% (dIII-E) and 4% (dIII-F) of the length of SV40 DNA. When a dIII digest of SV40 DNA was chilled at 0°C and examined by electron microscopy, a high percentage of the fragments were circular, thus confirming the presence of cohesive ends.

Digestion of each isolated dIII fragment with endo R·*Hin* showed the following: fragment dIII-A contains *Hin* fragments B, F, G, J; dIII-B corresponds to *Hin*-A; dIII-C contains *Hin*-C and D; dIII-D contains *Hin*-H and I; dIII-E corresponds to *Hin*-E; and dIII-F corresponds to *Hin*-K (Lee, Danna & Nathans, unpublished observations). From these results we can place these fragments in the previously determined *Hin* cleavage map, as shown in Figure 1.

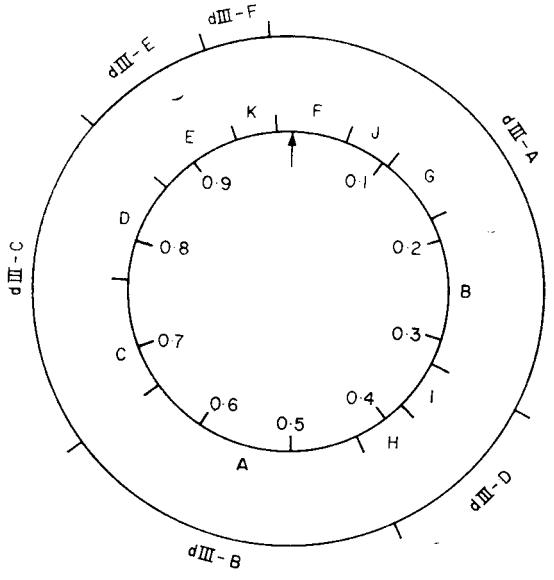

FIG. 1. *Hin* cleavage map of the SV40 genome (Danna *et al.*, 1973). The sites of cleavage by endo R·*Hin*dIII and resulting fragments are shown in the outer circle. Also shown is the *Eco* RI site (Morrow & Berg, 1972; Mulder & Delius, 1972) designated as the origin (arrow).

(c) *Preparation of deleted simian virus 40 DNA*

To construct deleted SV40 genomes, wild-type [^{14}C]DNA I (mixed with tracer [^{32}P]DNA I) was partially digested with endo R·*Hin*dIII, and the resulting fragments were subjected to electrophoresis in agarose gels followed by autoradiography. Reference to the map shown in Figure 1 indicates that the following large partial digest products would be expected: *Hin*-K deletion (96% of SV40 DNA length); *Hin*-E deletion (91·5%); *Hin*-HI deletion (89·5%); *Hin*-EK deletion (87·5%); *Hin*-CD deletion (79·5%); *Hin*-A deletion (77·5%); *Hin*-CDE deletion (71·5%), etc. As seen in Plate II, which shows the separation of the partial digest products, many fragments were in the size range of 100% to 50% of full-length SV40 DNA. Several discrete size classes of DNA were discernible in the autoradiogram, each of which was cut out of the gel and isolated by electrophoresis. Although most of these fractions were not homogeneous, they were nearly free of full-length molecules and therefore could be used for cloning by complementation plaqueing (see below).

(d) *Complementation plaqueing of deleted molecules*

To isolate clones of deletion mutants, excisionally deleted, linear molecules varying in length from about 96% to about 65% of SV40 DNA were used. (This size range

was chosen since it had been found earlier that defective SV40 virions formed during undiluted serial passage contained genomes varying from about 70% to approximately unit length (Brockman & Nathans, 1974); presumably this reflects a size limit of encapsidation of DNA.) BSC-40 cells growing in microwells were infected with a given size class of DNA fragments, followed by infection with either an early or a late ts mutant, as described in Materials and Methods. Infective centers were then plated at the restrictive temperature and allowed to form plaques. No plaques appeared when a ts mutant alone was used, and few plaques appeared with short DNA alone. However, co-infection with certain of the DNA fragments plus a ts mutant led to abundant plaque formation, generally five to ten times the number of plaques with DNA alone. Therefore, some of the fragments were able to complement ts mutants (Table 1).

TABLE 1

Frequency of short genomes from complementation plaques

Agarose band†	ts helper	No. of plaques tested	No. of plaques yielding 0, 1, 2, >2 size classes of short genomes‡			
			0	1	2	>2
3	A28	15	2	7	6	0
3	B4	3	0	2	1	0
4	A28	24	1	14	8	1
4	B4	16	1	12	3	0
5	A28	16	2	8	5	1
5	B4	0				
6	A28	6	0	2	3	1
6	B4	10	3	5	2	0
7	A28	3	0	1	2	0
7	B4	0				
8	A28	3	1	1	1	0
8	B4	2	1	0	1	0

† As numbered in Plate II.
‡ Determined by electrophoresis of viral DNA isolated from infected cells, as shown in Plate III.

(e) *Survey of plaques for deleted genomes*

To determine whether putative "complementation plaques" actually contained virus with deleted genomes, a large number of plaques were aspirated and used to prepare stocks by infecting BSC-1 cells as described previously (Brockman & Nathans, 1974). ^{32}P-labeled viral DNA was then prepared in BSC-1 cells by Hirt's procedure (Hirt, 1967), and was subjected to electrophoresis in agarose. As shown in Plate III and tabulated in Table 1, many of the plaques contained one or more electrophoretically discrete species of short genomes in addition to full-length helper genomes. Based on their mobility, which has been correlated with genome length (Brockman & Nathans, 1974), the short molecules vary from about 70% to about 90% of SV40 DNA. As will be shown below, the short molecules consist of specifically deleted SV40 DNA.

As shown in Plate III, some plaques yielded more than one size class of short DNA. In this case, a particular DNA species could be recovered and re-cloned by

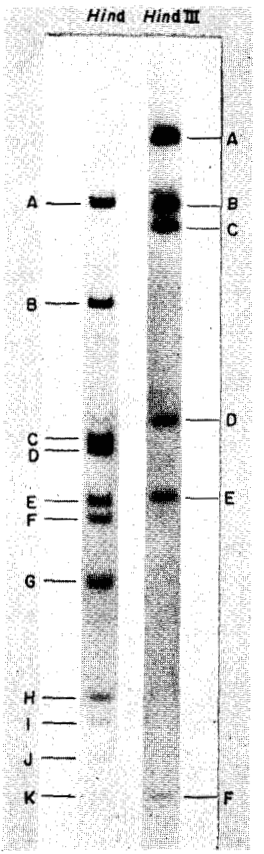

PLATE I. SV40 DNA fragments produced by endo R·*Hin*dIII. Autoradiogram of a polyacrylamide electropherogram of ^{32}P-labeled SV40 DNA digested with excess dIII enzyme and compared to a *Hin*d digest (Danna *et al.*, 1973). The origin is at the top. The positions of reference *Hin* fragments A to K are shown in the column marked *Hin*d. The actual distance between the origin and *Hin*-K is 30 cm.

[*facing p. 184*

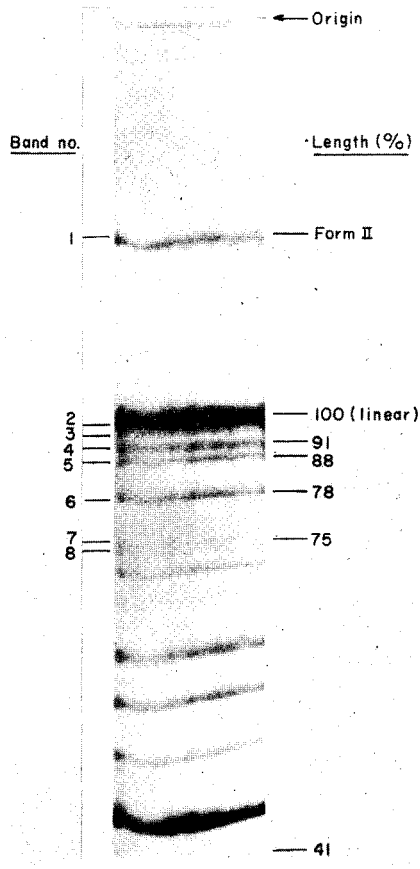

PLATE II. Partial digest products of SV40 DNA produced by endo R·*Hin*dIII. Autoradiogram of a preparative agarose electropherogram of ^{32}P-labeled SV40 DNA digested as described in Materials and Methods. On the scale at the right are the approximate lengths of fragments expressed as a percentage of SV40 DNA, based on correlation between length of linear DNA measured by microscopy and electrophoretic mobility. Conditions are those described previously (Brockman & Nathans, 1974).

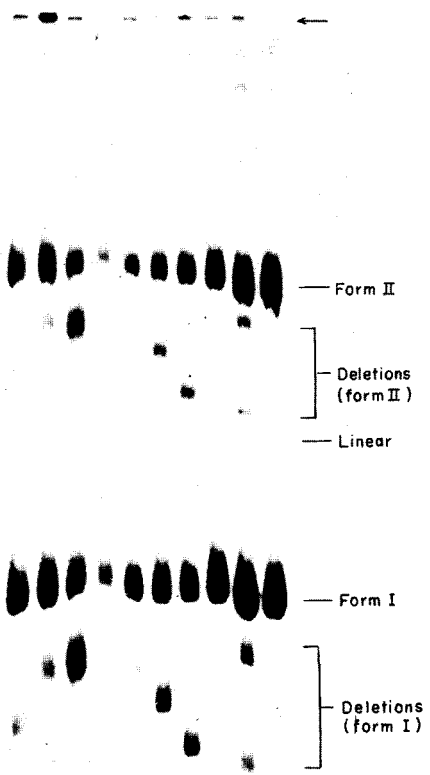

PLATE III. Survey of plaques for short genomes. ^{32}P-labeled viral DNA was isolated by Hirt's (1967) procedure from cells infected with virus stocks made from single plaques, and the DNA was subjected to electrophoresis in agarose gel slabs (Brockman & Nathans, 1974). A typical autoradiogram of the wet gel is shown. The positions of parental SV40 DNA forms I and II, linear SV40 DNA, and deleted DNA are indicated; the arrow indicates the origin.

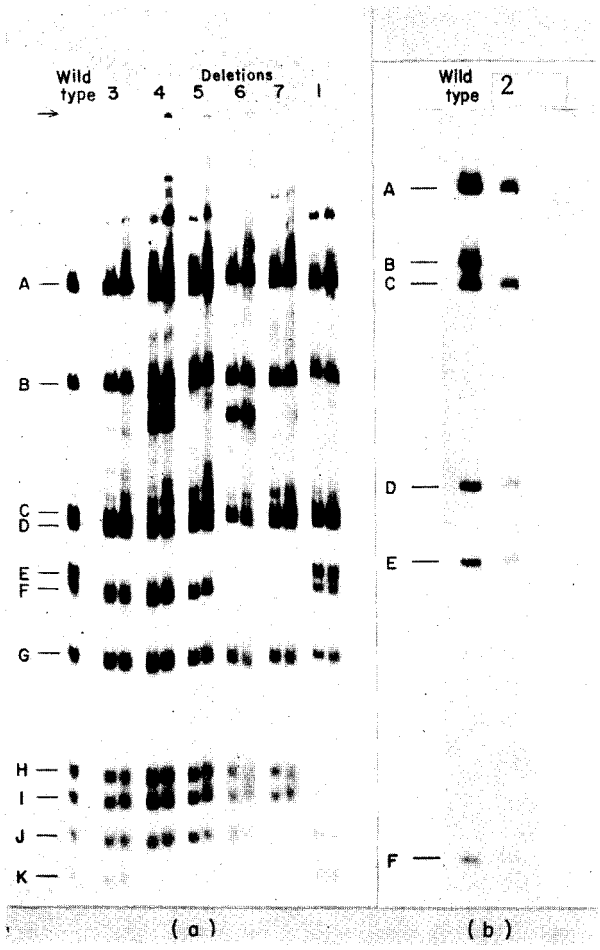

PLATE IV. *Hin* digests of deleted genomes. ^{32}P-labeled, deleted genomes recovered from agarose gels were digested with excess endo R·*Hin* (a) or *Hin*dIII (in the case of *dl*-1002, (b) and subjected to electrophoresis in polyacrylamide gels (Danna *et al.*, 1973) followed by autoradiography. A, B, C . . . K are the positions of parental *Hin* fragments (a) and A, B, C . . . F are parental dIII fragments (b) in the same gel slab. For each deleted genome (except *dl*-1002), two concentrations of enzyme were used to assess the extent of digestion. 1,2, etc. refer to *dl*-1001, *dl*-1002, etc.

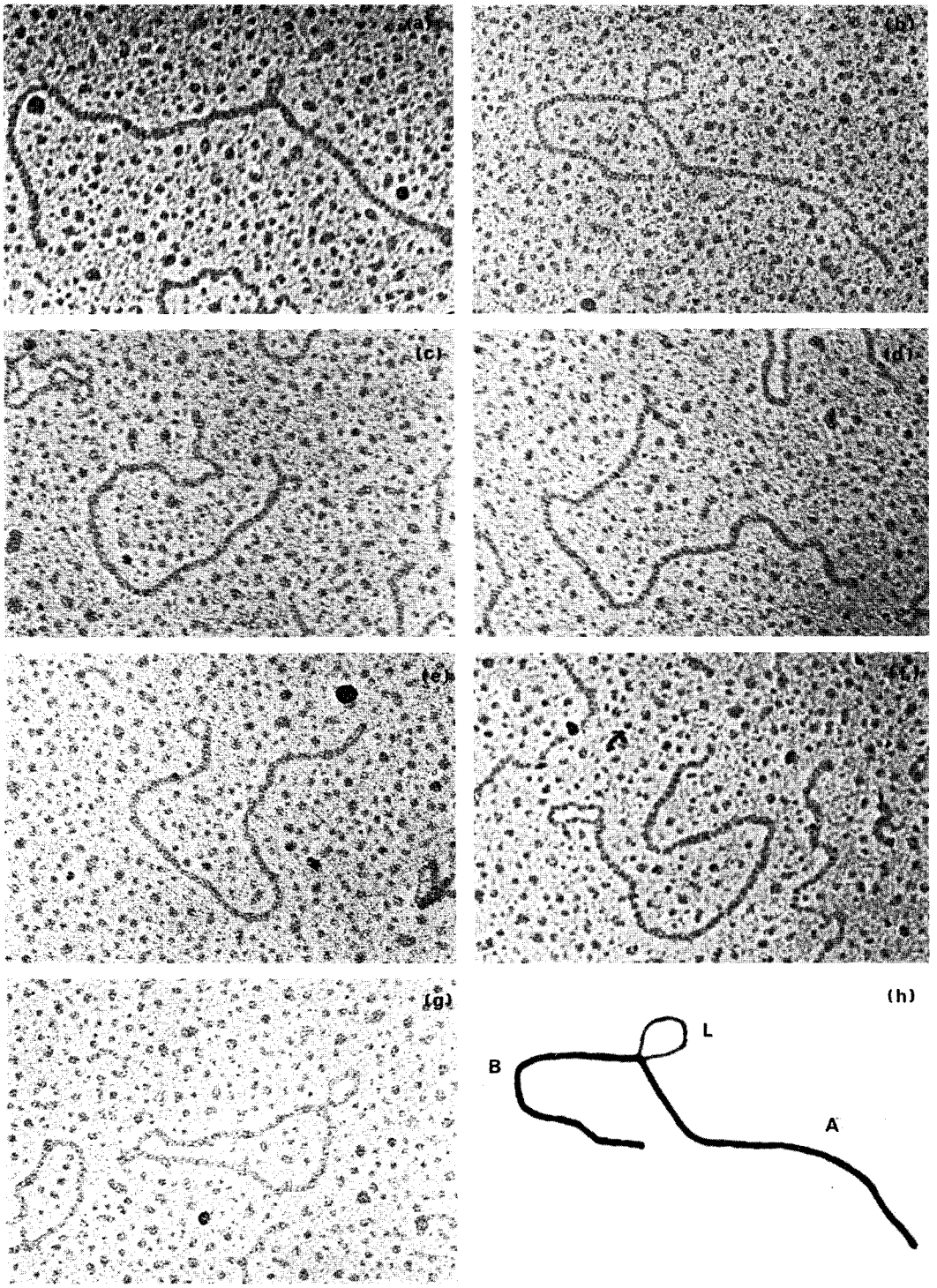

PLATE V. Heteroduplex molecules made between *Eco* RI linear molecules of deleted DNA and parental DNA. For each mutant, a typical heteroduplex molecule is shown in the micrograph. The letters in (h) refer to distances measured as given in Table 4.

(a) *dl*-1001–SV40 heteroduplex;
(b) *dl*-1002–SV40 heteroduplex;
(c) *dl*-1003–SV40 heteroduplex;
(d) *dl*-1004–SV40 heteroduplex;
(e) *dl*-1005–SV40 heteroduplex;
(f) *dl*-1006–SV40 heteroduplex;
(g) *dl*-1001–*dl*-1007 heteroduplex;
(h) diagram of *dl*-1002–SV40 heteroduplex.

complementation plaqueing, resulting in a deletion mutant with genome of uniform length. It should also be noted in Plate III that plaques which were produced by complementing short molecules of a given size class with a ts mutant yielded genomes of considerable size difference, including many that were shorter than expected. As will be shown, this heterogeneity is due to extension of the original excisional deletion to involve contiguous regions of the genome.

(f) *Endo R*·Hin *digests of short genomes*

To localize the site of deletion in each of the short genomes, *Hin* digests were compared to digests of wild-type, full-length DNA. For this purpose ^{32}P-labeled, deleted DNA was prepared from cells infected with virus stocks made from plaques containing a deletion mutant plus a helper ts mutant, and short viral DNA molecules were recovered from agarose gels. The DNA was then digested with endo R·*Hin*d (a combination of the dII and dIII enzymes), and the fragments were subjected to electrophoresis to determine which *Hin* fragments were present. A digest of an excisional deletion caused by the *Hin*dIII enzyme should be missing one or more contiguous *Hin* fragments, thus establishing the site of deletion. However, if the limits of the deletion do not coincide with *Hin* cleavage sites, a new fragment should also be present in the digest, representing fusion of the *Hin* fragments on either side of the deleted segment.

TABLE 2

Hin *fragments of excisional deletions*

Mutant	Agarose band (%)[a]	ts helper	Missing fragments	Mol. wt of fragments present (%)[b]
dl-1001[c]	4 (91)	B4	H, I	89·5
dl-1002[d]	6 (78)	B4	A	77·5
dl-1003[e]	4 (91)	A28	E	91·5
dl-1005[f]	5 (88)	A28	E, K	87·5

[a] Agarose band no. is designated in Plate II. The number in parentheses is the approximate length of the DNA molecules in the given band, as a percentage of SV40 DNA, based on electrophoretic mobility.

[b] Sum of the molecular weights of *Hin* fragments, as a percentage of SV40 DNA (Danna *et al.*, 1973).

[c] One of 6 independently cloned mutants lacking *Hin*-H and I only. All were derived from agarose band no. 4 and formed plaques with ts B4.

[d] One of 2 independently cloned mutants lacking *Hin*-A only. Both were derived from agarose band no. 6 and formed plaques with ts B4.

[e] One of 2 independently cloned mutants lacking *Hin*-E only. Both were derived from agarose band no. 4 and formed plaques with ts A28.

[f] One of 3 independently cloned mutants lacking *Hin*-E and K only. Two were derived from band no. 5 and one from band no. 4. All formed plaques with ts A28.

Examples of both types of deleted genomes are shown in Plate IV and the data for all the deletion mutants analyzed are tabulated in Tables 2 and 3. (In addition to those tabulated, several other mutants were detected, which probably arose by recombination between the dIII DNA fragment and DNA of the ts helper virus; see Discussion.) Deletion mutant *dl*-1001 yielded all *Hin* fragments except *Hin*-H and I and is therefore a *Hin*-HI deletion; *dl*-1002 is a *Hin*-A deletion; *dl*-1003 is a

TABLE 3

Hin *fragments of extended deletions*

Mutant	Agarose band (%)†	ts helper	Missing *Hin* fragments	Size of new fragment (%)§	Probable identity of new fragment¶	Mol. wt of fragments present (%)‡	Probable excisional deletion from which derived‖
dl-1004	4 (91)	A28	D,E,K	13·8	$\overline{DK}$	91·2	E^-
dl-1006	5 (88)	A28	D,E,F,K	13·7	$\overline{DF}$	83·7	$(E,K)^-$
dl-1007	5 (88)	A28	D,E,F,J,K	10·3	$\overline{DJ}$	75·8	$(E,K)^-$
dl-1008	4 (91)	A28	D,E,F,K	11·5	$\overline{DF}$	81·5	E^-
dl-1009	4 (91)	B4	A,B,H,I	27·0	$\overline{AB}$	79·0	$(H,I)^-$

†‡ See footnotes to Table 2.

§ Estimated by electrophoretic mobility and expressed as a percentage of SV40 DNA.

¶ These are denoted by the *Hin* fragments from which they are derived, e.g. the symbol $\overline{DK}$ denotes an *Hin* fragment with part of *Hin*-D at one end and part of *Hin*-K at the other end. Their identity was deduced from the *Hin* digest pattern and the *Hin* cleavage map (Fig. 1).

‖ These are denoted by the *Hin* fragment missing from the DNA, e.g. E^- is the linear dIII fragment lacking *Hin*-E, $(E,K)^-$ is the fragment lacking *Hin*-E and K, etc.

Hin-E deletion; *dl*-1005 is a *Hin*-EK deletion. As indicated by the cleavage map shown in Figure 1, E and K are contiguous fragments, as are H and I. Since none of these mutants yielded new fragments, they are all excisional deletions, i.e. the limits of the deletion correspond to dIII cleavage sites (see Fig. 1).

On the other hand, many short genomes were isolated, the digests of which lacked contiguous *Hin* fragments and also contained a single new fragment. Three examples are shown in Plate IV and others are presented in Table 3: *dl*-1004 lacks *Hin*-DEK, a continuous segment of SV40 DNA, and has a new fragment which is 13·8% of the length of SV40 DNA (Table 3). We presume that this fragment resulted from fusion of part of *Hin*-D and *Hin*-K (see Fig. 1). Similarly, *dl*-1006 (Plate IV and Table 3) lacks *Hin*-DEFK and has a new fragment which is 13·7% of unit length. This fragment is presumably a fusion of *Hin*-D and *Hin*-F segments (see Fig. 1). *dl*-1007 lacks *Hin*-DEFJK and has a new fragment which is 10·3% of unit length (between *Hin*-C and D). (The CD region is not well displayed in Plate IV, but is clear on the original film.) This new fragment is presumably a fusion of *Hin*-D and *Hin*-J segments (see Fig. 1). *dl*-1008 lacks *Hin*-DEFK and has a new fragment which is 11·5% of unit length; therefore it is similar, but not identical to *dl*-1006. *dl*-1009 lacks *Hin*-ABHI and contains a new fragment which is about 27% of unit length. Presumably this fragment resulted from fusion of *Hin*-A and *Hin*-B segments. It is noteworthy that except for *dl*-1004, each of the genomes which yielded a new *Hin* fragment has a length which is considerably shorter than the DNA from which it arose (see Table 3), in contrast to excisionally deleted genomes (Table 2). Based on the *Hin* digest pattern and the size of the original dIII fragment from which each mutant genome was generated, we have deduced the probable identity of the excisionally deleted linear fragment which gave rise to each of these mutants (Table 3, last column). Our general conclusion from the analysis of this group of mutants is that in addition to purely excisional deletions, other deletions (extended deletions) have arisen in which the original excisional deletion has been extended beyond the sites of enzymatic cleavage.

(g) *Heteroduplex mapping*

In order to confirm the map positions of excisional deletions and to localize more precisely the limits of extended deletions, we prepared heteroduplexes between linear deleted DNA and linear parental DNA made by cleavage with endo R·*Eco* RI (Mulder & Delius, 1972; Morrow & Berg, 1972). These were then examined in the electron microscope and individual molecules were measured. Plate V shows representative heteroduplex molecules from each of the seven deletion mutants examined, and Table 4 gives the results of measuring specific segments of each heteroduplex molecule.

As shown in Plate V, all the deleted molecules examined contained only a single deletion loop, as predicted from the *Hin* fragment patterns (Tables 2 and 3). The distance of each deleted segment from the RI cleavage site and the size of the deletion loop in each case were estimated by measuring individual molecules (Table 4). For example, *dl*-1001/SV40 heteroduplexes showed a deletion loop at 55·4 $\pm$ 1·8% (length A, Plate V) of the length of SV40 DNA from one end of the *Eco* RI linear molecule. Since only *Hin*-H and *Hin*-I are missing from the endo R·*Hin* digest (Plate IV), the deletion loop consists of the *Hin*-HI segment of the wild-type strand, and length A (the longer duplex arm) is the distance between the *Eco* RI site and the *Hin*-AH junction in the counterclockwise direction of the cleavage map (see Fig. 1),

TABLE 4

Measurements of heteroduplex molecules and map positions of deletions

Mutant	*Eco* RI site[a]	Lengths as a percentage of SV40 DNA[b] A[c]	B	L	Mutant DNA[d]	Sum of *Hin* fragments	Map positions of deletions From electron microscopy[e]	From endo R digest[f]
dl-1001	+	55·4±1·8	31·1±1·7	10·7±1·3	89·5±2·3	89·5	0·31—0·42	0·325—0·430
dl-1002	+	42·7±1·6	35·4±1·2	23·4±3·2	76·0±1·8	77·5	0·43—0·67	0·430—0·655
dl-1003	+	85·0±2·5	4·8±0·5	7·3±1·3	90·4±2·6	91·5	0·85—0·95	0·860—0·945
dl-1004	+	83·5±2·2	4·5±0·4	7·8±1·0	89·5±1·9	91·2	0·84—0·94	0·84—0·93[j]
dl-1005	+	83·2±3·3	—[h]	14·7±1·5	85·7±1·7	87·5	0·83—0·97	0·860—0·985
dl-1006	+	83·0±3·4	—[h]	18·0±1·8	81·7±2·1	83·7	0·83—1·00[i]	0·83—1·00
dl-1007[g]	—	41·3±1·4	27·0±1·4	27·2±5·3 10·4±1·1	75·7±1·9	75·8	0·83—0·07	0·83—0·07

[a] +, *Eco* RI site present; —, *Eco* RI site missing.

[b] The values are mean measured lengths ± S.D. as a percentage of reference SV40 DNA on the same grid.

[c] A, B, L are distances measured as defined in Plate V(h). For A and B the reference was SV40 DNA form II; for L the reference was single-stranded SV40 DNA.

[d] Form II mutant DNA was measured *versus* SV40 DNA-L_{RI} or *vice versa*.

[e] The first position is the distance from the *Eco* RI site to the deletion loop (A or B); the second position is the first position plus (1·00 minus the fractional length of mutant DNA).

[f] For excisional deletions, these values were taken directly from the cleavage map (Fig. 1). For extended deletions, the first position was taken from the electron microscope measurements, and the second position was determined from the molecular weight of *Hin* fragments.

[g] Since *dl*-1007 DNA lacks an RI site, it was hybridized to *dl*-1001 DNA (see Plate V). A and B are the distances between the two deletion loops; the smaller contains *Hin*-H,I.

[h] Length B was too short to measure (see Plate V).

[i] Although this end of *dl*-1006 is calculated as 0·01, the *Eco* RI site is present; therefore, the deletion does not extend beyond the origin of the map.

[j] These values were calculated as described in footnote f. However, since the *Hin* digest lacks *Hin*-D and *Hin*-K (Table 3) and contains a new fragment ($\overline{DK}$), which is only slightly smaller than *Hin*-D + *Hin*-K (13·8% *versus* 14·0%), the limits of the *dl*-1004 deletion are nearer to 0·859 to 0·946.

whereas length B (the shorter duplex arm) is the distance between the *Eco* RI site and the *Hin*-BI junction in the clockwise direction. The deletion loop (L) measured 11 $\pm$ 2% of the length of SV40 DNA, corresponding to the length of *Hin*-H plus *Hin*-I (Fig. 1). Therefore, we conclude that *dl*-1001 has a deletion extending from 0·32 to 0·43 map unit (Table 4). Each of the other heteroduplex molecules has been analyzed in a similar way (Table 4).

In one instance (*dl*-1007) the deletion included the *Eco* RI site. Therefore, this mutant was mapped by examination of heteroduplex molecules made between singly nicked, open circular DNA from *dl*-1007 and *dl*-1001 (HI deletion). The position of the HI deletion loop in this instance served as a reference point in the molecule (Plate V), allowing precise localization of the *dl*-1007 deletion (Table 4). Several other mutant DNA molecules were also annealed with *dl*-1001 DNA; in each case the distance between deletion loops was as predicted.

In the last two columns of Table 4 we have compared the results of heteroduplex mapping with the results of *Hin* digestion of each mutant DNA. In the case of excisional deletions, the *Hin* digests are sufficient to localize the deletion and provide a more precise estimate of its position and extent than does heteroduplex mapping,

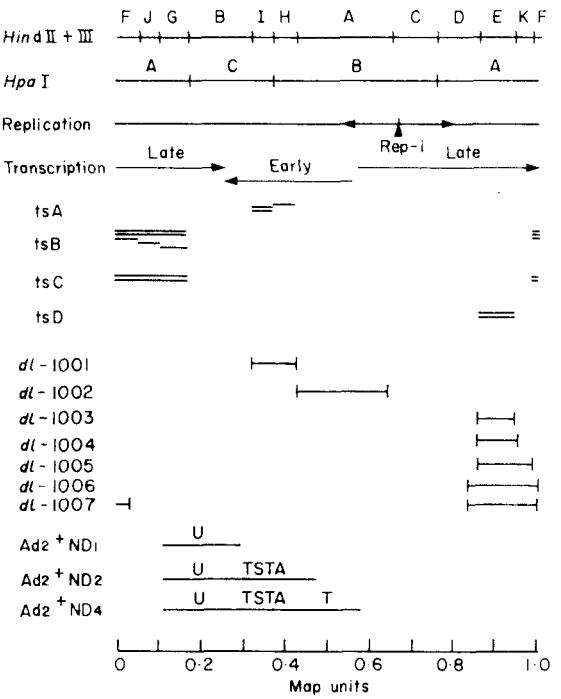

FIG. 2. A map of the SV40 genome indicating deletion sites of the mutants described in this paper, localization of ts mutations (Lai & Nathans, 1974), probable transcription sites and direction of transcription (Khoury *et al.*, 1973), initiation and termination sites for DNA replication (Danna & Nathans, 1972; Fareed *et al.*, 1972), and genetic determinants for certain early SV40 antigens (Kelly & Lewis, 1973; Morrow *et al.*, 1973; Lebowitz *et al.*, 1974). The circular map has been opened at the RI site. Rep-i is the origin of DNA replication; tsA, tsB, etc., refer to classes of ts mutants, and the bars indicate the positions of fragments which rescue mutants in a given class; U, TSTA, T refer to determinants of SV40-specific U antigen, tumor specific transplantation antigen, and tumor antigen, respectively, detected in cells infected with adeno–SV40 hybrid viruses (Ad_2+ND). tsA28 maps in *Hin*-I and tsB4 maps in *Hin*-J (Lai & Nathans, 1974).

owing to the sensitivity of electrophoretic mobility as a measure of DNA length (Danna *et al.*, 1973). In the case of extended deletions, the position of the deletion cannot be accurately determined by restriction enzyme digestion, but its length can be estimated more precisely than by heteroduplex mapping. Therefore the map positions given in the last column of Table 4 represent the best estimate of the localization and extent of each deletion.

(h) *A map of deletion sites*

A summary of the results for several of the mutants analyzed is shown in Figure 2 in the form of a map of the SV40 genome. The deletion site of each mutant is indicated in relation to the currently known loci of template functions and determinants of early antigens. In addition, the positions of recently mapped ts mutants are indicated in the Figure.

4. Discussion

We have described in this communication a method for generating deletion mutants of SV40 which depends on the enzymatic excision of specific segments from the viral DNA and subsequent cloning of deleted genomes by complementation with ts mutants (Brockman & Nathans, 1974). The deletion mutants were mapped by analysis of endo R·*Hin*d digests in comparison with the *Hin* cleavage map, and by heteroduplex mapping. Among the different mutants reported, two have deletions within the early region of the genome, as defined by transcription mapping, six are within the late region, and one involves both regions (Fig. 2 and Table 4).

Reference to the cleavage map shown in Figure 1 indicates that it should be possible to construct six deleted genomes of size appropriate for encapsidation, by excision of *Hin*dIII fragments. Of these, we have cloned all but the dIII-C (*Hin*-CD) deletion and the dIII-F (*Hin*-K) deletion. The genome lacking *Hin*-K (96% of SV40 DNA) may not have been well separated from full-length helper virus genomes and therefore not recognized (see Plate III). However, the genome lacking *Hin*-CD (79% of SV40 DNA) is probably missing among the cloned deletion mutants because it lacks the initiation site for DNA replication present in *Hin*-C (Danna & Nathans, 1972). Such a site presumably would be required for replication of any variant genome.

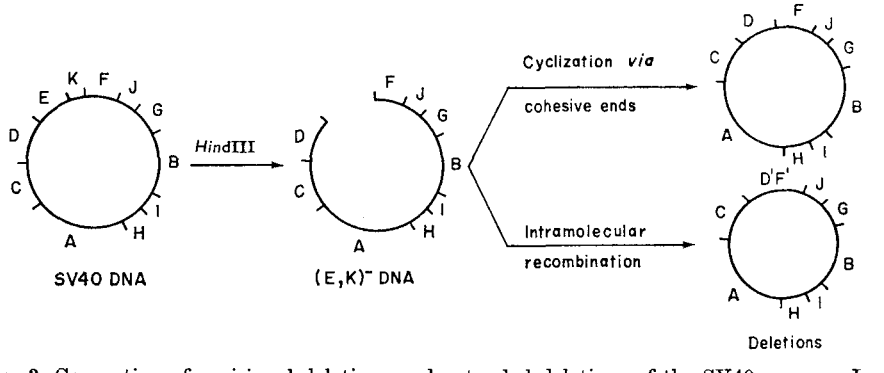

FIG. 3. Generation of excisional deletions and extended deletions of the SV40 genome. In the illustrations, excision of *Hin*-E,K by endo R·*Hin*dIII leads to linear (E,K)⁻ DNA with cohesive ends. Cyclization *via* cohesive ends leads to (E,K)⁻ circular genomes, and intramolecular recombination leads to extension of the original excisional deletion and cyclization.

The mutants analyzed in this report were all derived from partial digest products of SV40 DNA made with endo R·*Hin*dIII, since we originally expected a requirement for fragments with cohesive ends in order to effect circularization (see Fig. 3). However, in addition to the anticipated mutants, which lacked specific segments of DNA corresponding to regions between dIII sites, several deletion mutants were found in which the deletion extended beyond the cleavage sites. These genomes must have been generated within the cell by recombination, presumably intramolecular recombination near the ends of fragments. Although the enzymatic steps needed for this type of recombinational event are not known, a plausible mechanism involves sequential exonuclease action, base-pairings, and repair, resulting in covalently closed circles with extended deletions (Fig. 4). Since each of the extended deletions so far analyzed was distinctive (Table 3), it is likely that this type of recombination can occur at many sites in the SV40 genome.

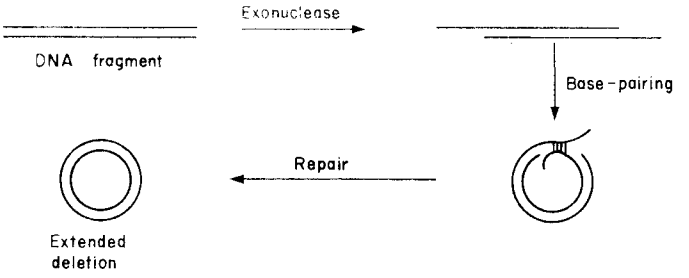

FIG. 4. A hypothetical mechanism for intramolecular recombination near the termini of DNA fragments.

Aside from mutants which probably arose by intramolecular recombination, several mutants were isolated which appear to have arisen by recombination between the DNA of a fragment and the DNA of a ts helper virus (Lai & Nathans, unpublished observations). Since the *Hin* digest of DNA from the ts helper differs from that of DNA from which the deleted genomes were derived (Lai & Nathans, 1974), recombinants could be identified by their distinctive *Hin* digests. These will be described in a later report.

A practical consequence of the intramolecular recombinational generation of deletions is the possibility of isolating a great variety of deletion mutants by using site-specific or non-specific endonucleases, whether they produce "staggered" or even-ended breaks in DNA. Even linear, full-length SV40 DNA containing non-cohesive termini (*Hpa*I linear molecules) has yielded cloned populations of deletion mutants by complementation plaqueing (Lai & Nathans, unpublished observations).

As pointed out earlier, mapped deletion mutants of SV40 should be of value in genetic and physiological studies. Excisional deletions will be especially useful. In contrast to deletions cloned from high passage virus stocks (Brockman & Nathans, 1974), their genomes are originally constructed *in vitro* from wild-type DNA, and are never exposed to high selective pressure. Neither are they exposed to mutagens as in the usual procedures for producing deletions. Therefore they are unlikely to have multiple alterations in the same molecule.

Of more general import is the possible extension of the methodology described in this communication to construction and cloning of entirely novel viral genomes.

Enough is now known of the organization of the SV40 genome (see Fig. 2) to allow rational enzymatic re-structuring of SV40 DNA for specific purposes. Such genomes can be cloned selectively by complementation plaqueing or non-selectively by co-infection with wild-type virus (Brockman & Nathans, 1974). Related studies of this type have recently been reported with bacterial R factor DNA (Cohen *et al.*, 1973).

Note added in proof: Since submission of this manuscript, two related studies have come to our attention: (1) Herzberg *et al.* (personal communication) have isolated deletion mutants of SV40 using endo R · *Eco*RII; and (2) N. Murray & K. Murray (*Nature (London)* 1974, in the press) have constructed excisional deletion mutants of phage λ using endo R · *Eco*RI and found extensions of the deletions beyond the RI cleavage sites.

We thank T. J. Kelly Jr for continuous guidance in electron microscopy and heteroduplex mapping procedures, T. N. H. Lee and W. W. Brockman for valuable help and advice, and Douglas Welsh and Jack DeBoy for able technical assistance. This research was supported by grants from the National Cancer Institute (CA-11895-03), the American Cancer Society, and The Whitehall Foundation.

This is publication no. 11 in a series on the genome of simian virus 40. Publication no. 10 is Lai & Nathans (1974).

REFERENCES

Brockman, W. W. & Nathans, D. (1974). *Proc. Nat. Acad. Sci., U.S.A.* **71**, 942–946.
Brockman, W. W., Lee, T. N. H. & Nathans, D. (1973). *Virology*, **54**, 384–397.
Chou, J. Y. & Martin, R. G. (1974). *J. Virol.* **13**, 1101–1109.
Cohen, S., Chang, A. C. Y., Boyer, H. W. & Helling, R. B. (1973). *Proc. Nat. Acad. Sci., U.S.A.* **70**, 3240–3244.
Danna, K. J. & Nathans, D. (1971). *Proc. Nat. Acad. Sci., U.S.A.* **68**, 2913–2917.
Danna, K. J. & Nathans, D. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3097–3100.
Danna, K. J., Sack, G. H., Jr & Nathans, D. (1973). *J. Mol. Biol.* **78**, 363–376.
Davis, R. W., Simon, M. & Davidson, N. (1971). In *Methods in Enzymology* (Grossman, L. & Moldave, K., eds), vol. 21, pp. 413–428, Academic Press, New York.
Dubbs, D. R., Rachmeler, M. & Kit, S. (1974). *Virology*, **57**, 161–174.
Fareed, G. C., Garon, C. F. & Salzman, N. P. (1972). *J. Virol.* **10**, 484–491.
Hedgpeth, J., Goodman, H. M. & Boyer, H. W. (1972). *Proc. Nat. Sci., U.S.A.* **69**, 3448–3452.
Hirt, B. (1967). *J. Mol. Biol.* **26**, 365–369.
Kelly, T. J., Jr & Lewis, A. M., Jr (1973). *J. Virol.* **12**, 643–652.
Khoury, G., Byrne, J. C. & Martin, M. A. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 1925–1928.
Khoury, G., Martin, M. A., Lee, T. N. H., Danna, K. J. & Nathans, D. (1973). *J. Mol. Biol.* **78**, 377–389.
Khoury, G., Fareed, G. C., Berry, K., Martin, M. A., Lee, T. N. H. & Nathans, D. (1974). *J. Mol. Biol.* **87**, 289–302.
Kimura, G. & Dulbecco, R. (1972). *Virology*, **49**, 394–403.
Kimura, G. & Dulbecco, R. (1973). *Virology*, **52**, 529–534.
Lai, C.-J. & Nathans, D. (1974). *Virology*, in the press.
Lavi, S. & Winocour, E. (1972). *J. Virol.* **9**, 309–316.
Lebowitz, P., Kelly, T. J., Jr, Nathans, D., Lee, T. N. H. & Lewis, A. M., Jr (1974). *Proc. Nat. Acad. Sci., U.S.A.* **71**, 441–445.
Martin, M. A., Gelb, L. D., Garon, C., Lee, T. N. H., Sack, G. H., Jr & Nathans, D. (1974). *Virology*, **59**, 179–189.
McCutchan, J. H. & Pagano, J. S. (1968). *J. Nat. Cancer Inst.* **41**, 351–357.
Mertz, J. E. & Davis, R. W. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3370–3374.
Morrow, J. F. & Berg, P. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3365–3369.
Morrow, J. F., Berg, P., Kelly, T. J., Jr & Lewis, A. M., Jr (1973). *J. Virol.* **12**, 653–658.

Mulder, C. & Delius, H. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3215–3219.
Robb, J. A., Tegtmeyer, P., Ishikawa, A. & Ozer, H. A. (1974). *J. Virol.* **13**, 662–665.
Roy, P. H. & Smith, H. O. (1973). *J. Mol. Biol.* **81**, 427–459.
Rozenblatt, S., Lavi, S., Singer, M. F. & Winocour, E. (1973). *J. Virol.* **12**, 501–510.
Sambrook, J., Sharp, P. A. & Keller, W. (1972). *J. Mol. Biol.* **70**, 57–71.
Sambrook, J., Sugden, B., Keller, W. & Sharp, P. A. (1973). *Proc. Nat. Acad. Sci., U.S.A.* **70**, 3711–3715.
Smith, H. O. (1974). In *Methods in Molecular Biology Series* (Laskin/Last), vol. 9, Marcel Dekker, Inc., New York, in the press.
Smith, H. O. & Wilcox, K. W. (1970). *J. Mol. Biol.* **51**, 379–391.
Tai, H. T., Smith, C. A., Sharp, P. A. & Vinograd, J. (1972). *J. Virol.* **9**, 317–325.
Tegtmeyer, P. (1972). *J. Virol.* **10**, 591–598.
Tegtmeyer, P. & Ozer, H. L. (1971). *J. Virol.* **8**, 516–524.
Yoshiike, K. (1968). *Virology*, **34**, 391–401.